CDF/ANA/TOP/CDFR/CDF10635
August , 2011

Measurement of the Top Quark Mass in the Di-lepton Channel Using the Dalitz-Goldstein Method

The CDF Collaboration

www-cdf.fnal.gov

Abstract

This is a measurement of the mass of the top quark using a method developed by G. Goldstein and R.H. Dalitz. It is based on 2.0 fb^{-1} of data collected by the Collider Detector Facility at Fermi National Accelerator Laboratories. Di-lepton events were observed from colliding protons with anti-protons with $\sqrt{s} = 1.96$ TeV in the Tevatron Collider. A total of 145 candidate events were observed with 49 expected to be from background when no b-Tagging was used. (64 candidate events were observed with 5 expected to be from background when b-Tagging was used.) These events include two neutrinos which elude detection. The method solves for the two neutrino momenta using a geometrical construction and taking as input the measured charged lepton and jet 4-momenta. In our analysis, we sample over a range of likely top quark masses choosing the most consistent mass with the help of an appropriately defined likelihood function. An important distinguishing feature of this method from others is its lack of dependence on the missing transverse energy, a quantity that is poorly measured by the experiment. This analysis determines the top quark mass to be $M_{top} = 172.3 \pm 3.4(stat.) \pm 2.1(syst.)$ GeV/c^2 ($M_{top} = 170.5 \pm 3.7(stat.) \pm 1.7(syst.)$ GeV/c^2 with b-tagging).

The Dalitz and Goldstein Method

1.1 Introduction

This paper will give an exposition of the Dalitz and Goldstein method[1] of measuring the mass of the top quark in the di-lepton channel. It is based on a geometrical construction and the leading order matrix elements for $t\bar{t}$ production, $t \to Wb$ decay, and $W \to l\nu$ decay, which defines a likelihood. The method has been modified to also work in the lepton plus jets channel by Dalitz, Goldstein, and Sliwa[2][3][4][5]. A major advantage of this method is the inclusion of both $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$ matrix elements and its lack of dependence on $\not{E}_T$, a point that is stressed by Sliwa in reference[6].

The original method was improved significantly in 1999 by Karr and Sliwa[7]. This analysis employs a new, much improved version of the technique, a result of careful studies performed over the past few years.

1.2 Geometrical Construction

The Dalitz and Goldstein[1] method for measuring the top quark mass in the di-lepton channel employs a geometrical interpretation of the equations of constraint.
Those equations are:

$$
\begin{aligned}
(l^+ + \nu)^2 &= M_W^2 && (1)\\
(l^- + \bar{\nu})^2 &= M_W^2 && (2)\\
(t - l^+ - b)^2 &= M_\nu^2 = 0 && (3)\\
(\bar{t} - l^- - \bar{b})^2 &= M_\nu^2 = 0 && (4)\\
t^2 = M_t^2 &= M_{\bar{t}}^2 = \bar{t}^2 && (5)\\
-P_x^t &\sim P_x^{\bar{t}} && (6)\\
-P_y^t &\sim P_y^{\bar{t}} && (7)
\end{aligned}
$$

where t, l, b, and ν are the top quark, lepton, bottom quark, and neutrino 4-momenta. M_W, M_t, and M_ν are the masses of the W boson, top quark, and neutrino. P_x^t and P_y^t are the x and y components of the top and anti-top quarks' transverse momenta. Equations 6 and 7 are only approximate and are considered "weak" constraints while the rest are "hard" constraints. Equations 6 and 7 are "weak" due to the possibility that the partons inside of the protons may have some transverse momentum. If this were not the case, then both equations 6 and 7 would be "hard" constraints.

The geometrical construction begins by rewriting equations 1 through 4 in terms of the top and bottom quark kinematics

$$
\begin{aligned}
(\vec{P}_t - \vec{P}_b)^2 &= (E_t - E_b)^2 - M_W^2 \equiv R_W^2, && (8)\\
(\vec{P}_t - \vec{P}_b - \vec{P}_{l^+})^2 &= (E_t - E_b - E_{l^+})^2 \equiv R_\nu^2. && (9)
\end{aligned}
$$

$\vec{P}_t$, $\vec{P}_b$, and $\vec{P}_{l^+}$ are the 3-momenta for the Top, Bottom, and charged lepton; and E_t, E_b, and E_{l^+} are their corresponding energies. A second pair of equations can be written for the anti-top quark and its decay products in exactly the same way. These equations are the equation for a sphere in the 3-dimensional momentum space.

$$X^2 + Y^2 = R^2 \tag{10}$$

The centers of these two spheres are separated by the charged lepton's 3-momentum $\vec{P}_{l^+}$. Figure 1 shows an example of two non-intersecting spheres that have their centers separated by $\vec{P}_{l^+}$. Both $\vec{P}_{l^+}$ and $\vec{P}_b$ in this picture are fixed because they are measured quantities; however, $\vec{P}_t$ is not measured and free to move about. Since the two spheres in Figure 1 do not intersect, it is not possible to draw a $\vec{P}_t$ vector that satisfies Eq. 8 and 9 simultaneously. Figure 2 shows two spheres that intersect at a single point. In this case, there exists one $\vec{P}_t$ vector that will satisfy both equations. This is a special case that will be discussed more in the next paragraph. Figure 3 shows a more generic case where the two spheres intersect at more than one point. In this case, the intersection is a circle. As long as $\vec{P}_t$ lies on this circle, it will satisfy both equations.

If the lepton masses are neglected, the top quark energy will be constant on this circle with radius r (Figure 4).

$$r^2 = \frac{M_W^2}{\mid \vec{P}_{l^+} \mid}(E_t - E_o) \tag{11}$$

$$E_o = E_b - E_l + \frac{M_W^2}{4E_l} \tag{12}$$

E_0 is the lowest possible energy for the top quark given the kinematics of the event. If E_t is equal to E_0, this corresponds to the special case mentioned above. If E_t is less than to E_0, the two spheres will not intersect and therefore give no solution.

Given a different top quark energy E_t, a different pair of spheres can be constructed. For each pair of spheres that intersect, a new circle is created. The top quark energy will increase in the direction of $\vec{P}_{l^+}$. These circles form the surface of a paraboloid (Figure 5). While the top quark energy is constant on these circles, the top quark mass , M_t, is not. However, if the mass is fixed the top quark 3-momentum vector will be confined to a conic section of the paraboloid which happens to be, out of the four possible conic section types, an ellipse. The orientation and eccentricity of the ellipse will depend on the assumed mass and the 4-momenta of the leptons and Bottom Quarks (Figure 6). A similar ellipse can be constructed in the same way for the anti-top quark.

Once both ellipses are constructed, they are projected into the transverse plane of the detector, one of which is reflected about the origin. Using the "weak" constraints, Eq. 6 and Eq. 7, each pair of points, one from each ellipse, will correspond to a possible solution consistent with the assumed top mass and the measured lepton and jet momenta. If these final

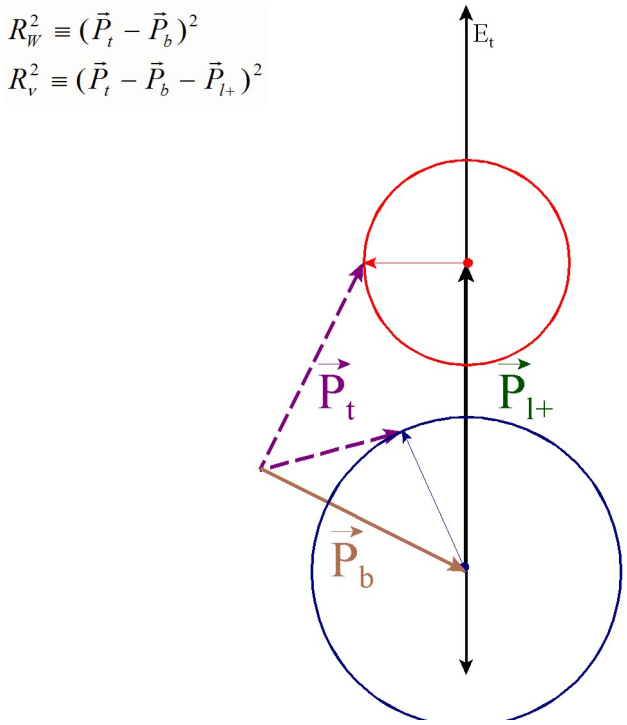

Figure 1: The two spheres of radii R_W and R_ν, whose centers are separated by $\vec{P}_t$, but do not intersect.

constraints were "hard", then only the points of intersection of the two ellipses would need to be considered as possible solutions. It is because they are "weak" that every pair of points needs to be considered. Given a pair of points from the projection of the two ellipses, the transverse momentum, $Pt_{t\bar{t}}$, of the top-anti-top system can be calculated. Each pair of points is assigned a likelihood factor, $P(Pt_{t\bar{t}})$, from the $Pt_{t\bar{t}}$ spectrum. The expected shape of the $Pt_{t\bar{t}}$ distribution is calculated from Monte Carlo simulation. If a different top mass is assumed, a different pair of ellipses will be created whose projections into the x-y plane will give another set of possible solutions.

1.3 The Likelihood

To determine the most likely mass of the top quark, a probability distribution is constructed for each combination of leptons and jets in an event. Given an assumed top quark mass, a likelihood value, L_i, for all pairs of points on the families of ellipses are projected onto the M_t-axis of this distribution. The most likely top quark mass corresponds to the peak of this distribution. The likelihood values, L_i, are a product of six probability factors.

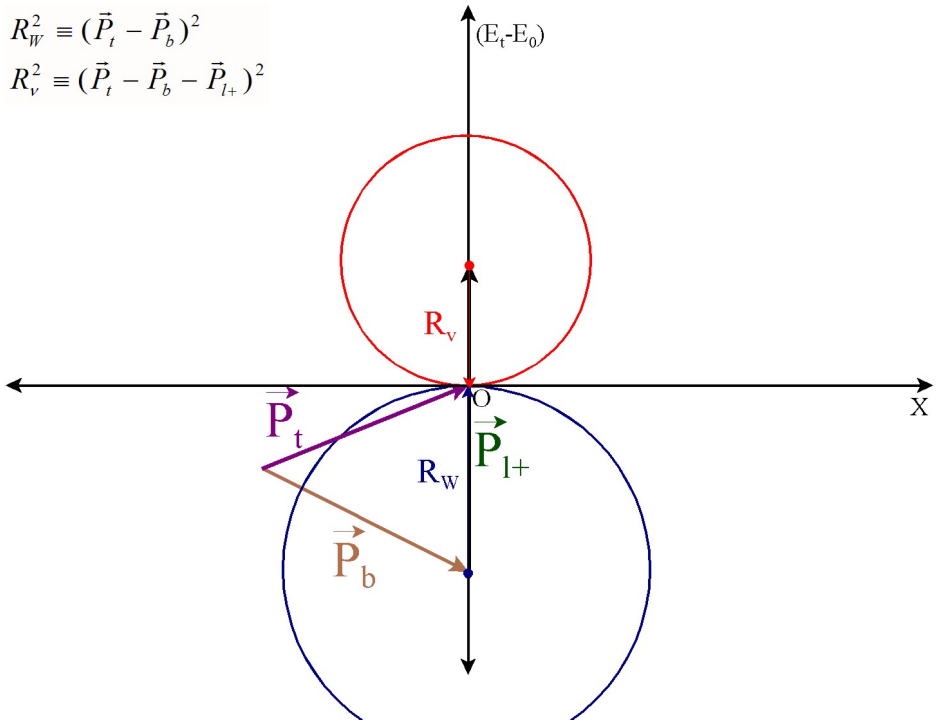

Figure 2: The two spheres of radii R_W and R_ν, which intersect at a point. O is the point where $E_t - E_0 = 0$, the minimum energy allowed for the top quark as determined by the kinematics of the event.

$$L_i \quad = \quad P(P_{t\bar{t}}) \times G(b) \times G(\bar{b}) \times P(x_1, x_2) \times P(l^+) \times P(l^-) \tag{13}$$

$P(Pt_{t\bar{t}})$ is the factor related to the transverse momentum of the top-anti-top system
$G(b)$ and $G(\bar{b})$ are the factors related to Jet Energy Smearing
$P(x_1, x_2)$ is the factor related to the leading order matrix element for the $t\bar{t}$ production.
$P(l^+)$ and $P(l^-)$ are the factors resulting from V-A Calculations with the matrix elements for the $t \to Wb$ and $W \to l\nu$ decays

i. $P(Pt_{t\bar{t}})$

Given a pair of points from the projection of the two ellipses, the transverse momentum, $Pt_{t\bar{t}}$, of the top-anti-top system can be calculated. Each pair of points is weighted by a likelihood factor, $P(Pt_{t\bar{t}})$, from the $Pt_{t\bar{t}}$ spectrum. The expected shape of the $Pt_{t\bar{t}}$ distribution is calculated from Monte Carlo simulation. If a different top mass is assumed, a different pair of ellipses will be created whose projections into the x-y plane will give another set of possible solutions.

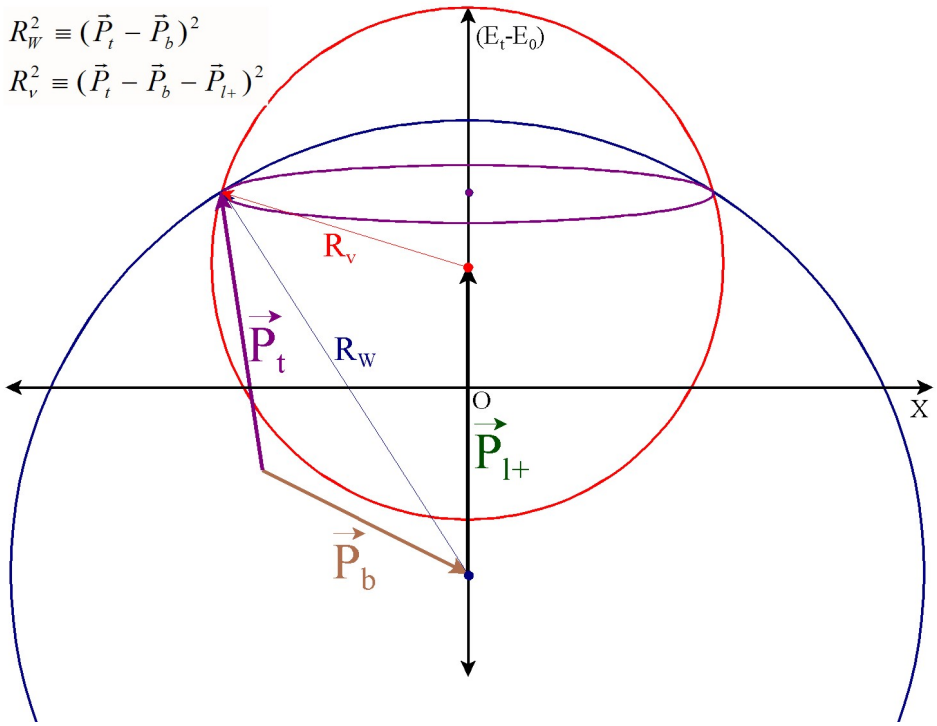

Figure 3: The two spheres of radii R_W and R_ν, showing a circle of intersection.

ii. $P(x_1, x_2)$

A relative likelihood factor is assigned that describes the level of agreement between the Feynman-x values, x_1 and x_2 , that are calculated from the event and those predicted by theory, i.e. the dependence of the leading order matrix element on the structure functions for the event,

$$P_{x_1,x_2} = \frac{\sum_{i=qq,gg} F_i(x_1)F_i(x_2)\frac{d\sigma}{d\hat{t}}(\hat{s},\hat{t})_i}{\sum_{i=qq,gg}\frac{d\sigma}{d\hat{t}}(\hat{s},\hat{t})_i} \quad (14)$$

$$x_{1,2} = (E_t + E_{\bar{t}} \pm (t_L + \bar{t}_L))/2P \quad (15)$$

$$\hat{s} = x_1 x_2 s \quad (16)$$

$$\hat{t} = M_t^2 - x_1\sqrt{s}(E_t - t_L) \quad (17)$$

where i labels the $q\bar{q}$ and gg processes; F_i are the structure functions; $\hat{s}$ is the parton-parton center-of-mass energy; $\hat{t}$ is the momentum transfer of the top-anti-top quark production subprocess; P is the proton momentum; s is the square of the energy in the proton-anti-proton system; and t_L is the longitudinal momentum of the top quark in the lab frame of the proton-anti-proton system.

iii. $P(l^+)$ and $P(l^-)$

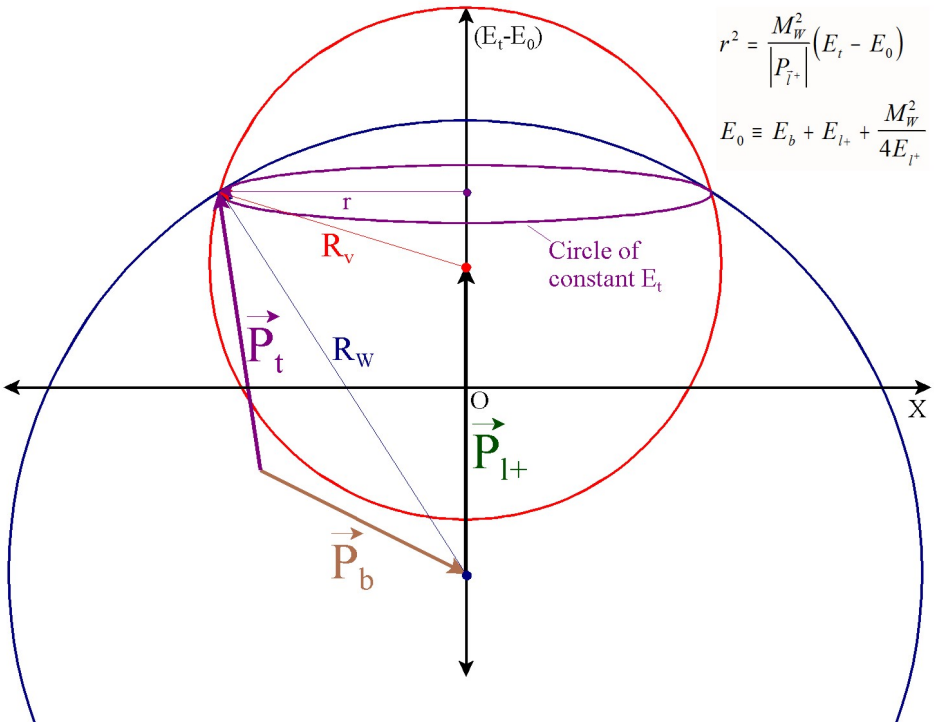

Figure 4: The two spheres of radii R_W and R_ν showing a circle of intersection with constant energy E_t.

For each point a likelihood factor is calculated to describe the agreement of the charged lepton energies calculated in the top quark rest frame with the values predicted by V-A calculations of the $t \to W + b$ and $W \to l + \nu$ decays.

$$dP(E_l) \quad = \quad (24/M_t^2)E_l(1 - 2E_l/M_t)dE_l \tag{18}$$

iv. $G(b)$ and $G(\bar{b})$

Since the b-jet and b-quark measured energies have large errors, the true parton energies can differ from the measured values. A range of energies is defined which is centered about the measured energy of each of the jets in an event. The range is chosen to be 3σ, where σ is the width of the jet energy resolution distribution. A finite number of jet energy points is chosen within the 3 signa range, and a probability, G(b), is assigned to each point. Points that correspond to b-jet energies that deviate from the measured value will be downgraded by a Gaussian probability factor giving them a lower probability than ones closer to the measured value. This smearing of the jet energies will give a family of ellipses for each jet. Each combination from an event with different jets will have its own pair of sigmas, one for each jet in the combination. The number of discrete points is set to the same value for all jets even though their sigmas will vary. In order to compensate for this, the likelihoods are first multiplied by the product of these sigmas when comparing the likelihood of one combination to another.

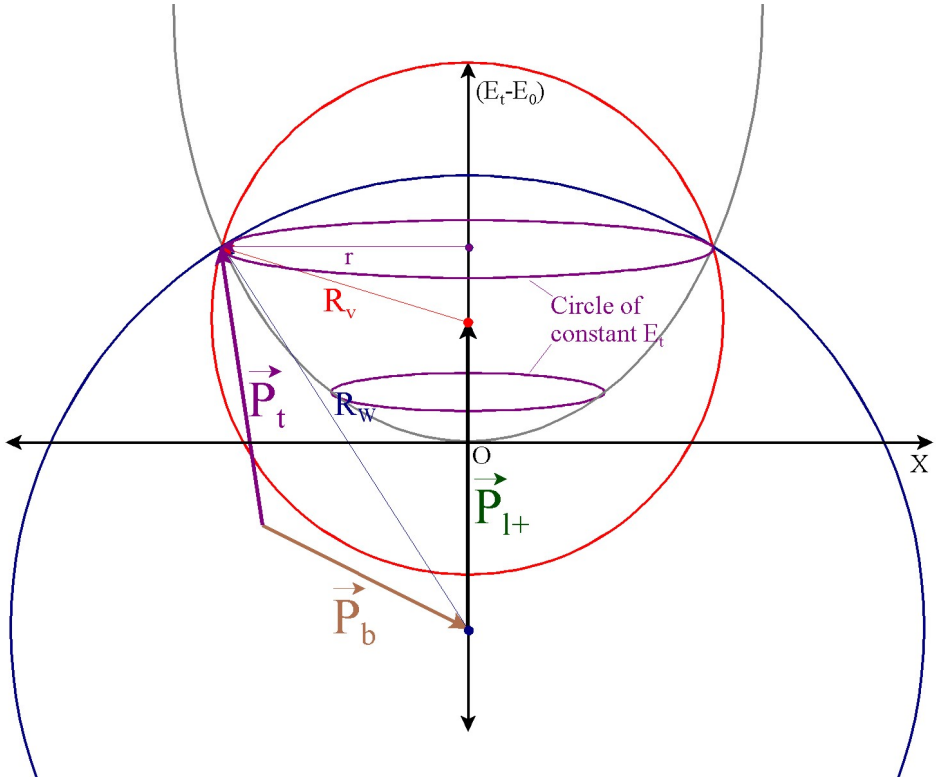

Figure 5: The two spheres of radii R_W and R_ν showing a circle of intersection with constant energy E_t. Also seen here is a second circle of constant E_t resulting from the intersection of two other spheres (not drawn). Both circles begin to form the paraboloid surface.

Pairing together all the combinations of ellipses from each family will create a 2-dimensional grid where the axes are labeled by the indeces of the smeared jet energy points. Each point on the grid will have a value which is the L_i that corresponds to smeared jets momenta, the given lepton momenta, and assumed top mass. There will be a different grid for each assumed top mass. Altogether, this will create a 3-dimensional space where the third axis is the assumed top mass. Each point in this space will have a value that comes from the sum of all the pairs of points on the pair of ellipses that correspond to the given assumed top mass and the two smeared jet energy values. The ellipses will vary in size. Geometrically speaking, each ellipse has an infinite number of points, but in the world of computers an infinite number of points can not be summed. Thus, a discrete set of points needs to be defined for each ellipse. This can be done in two ways. The first is to choose the number of points for all ellipses and allow the distance between points to be variable. The second way is to hold the distance between each point on an ellipse constant, so the number of points in the sum will be its circumference divided by the pre-defined distance between points. The first way has a significant problem. The most important points in the sum over the pairs of points on the ellipses are the intersection points of the two ellipses. If the distance between points is allowed to be variable, and an ellipse is large, it becomes more likely that the intersection point and the points around the intersections

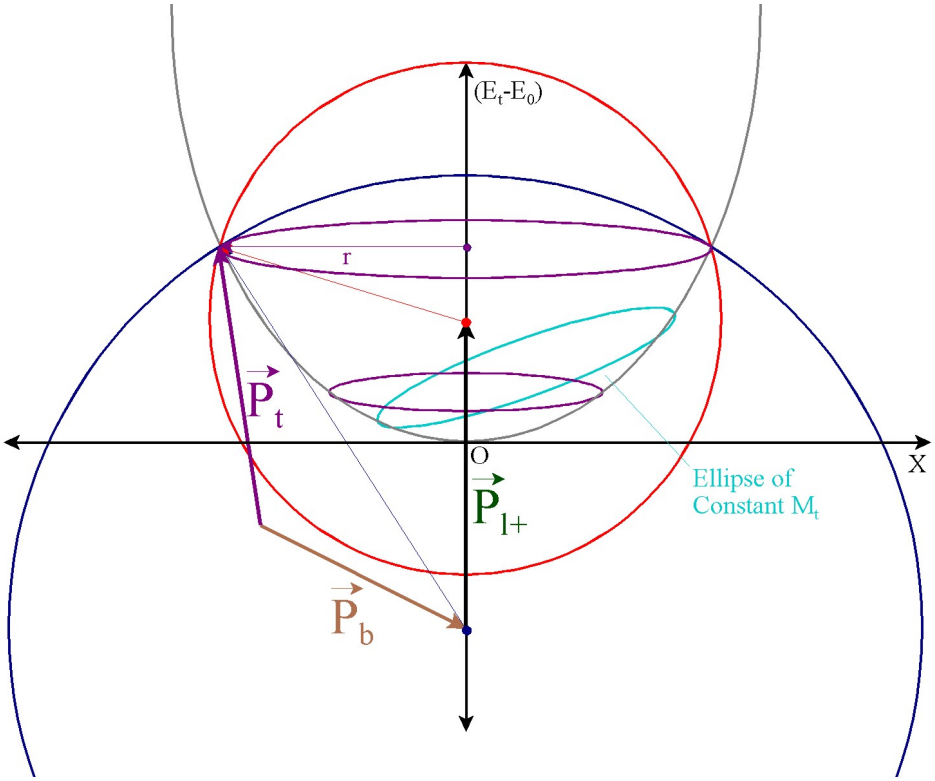

Figure 6: The two spheres of radii R_W and R_ν showing a circle of intersection with constant energy E_t. Assuming a constant M_t confines P_t to a conic section of the paraboloid which is an ellipse.

will be missed. The advantage here is that the number of points that are being considered can be kept small and thus reduces the amount of CPU time needed to do the analysis. This was the method used by Karr and Sliwa[7] in their measurement in 1999. The second method has two possible drawbacks. If an ellipse is small, it will have very few points, possibly as little as one point, and if an ellipse if very large, it will take a significant amount of CPU time to compute the sum. However, there should only be small ellipses around the minimum allowed assumed top mass, and modern computers can compute these sums in a reasonable amount of time. The second method is superior in that it will guarantee integration near the intersection points given a small enough choice of the spacing. As newer and faster computers are created, a smaller spacing can be used which should help to improve the integration over the ellipses and thus improve this method of measuring the top quark mass. This is one of the important changes made in this measurement from the earlier measurement. Summing over the assumed top mass axis of this space will reduce it back to a 2-dimensional grid with the smeared jet energies as the remaining axes; however, now the value at each point is a total likelihood. The point on the grid with the greatest total likelihood will be chosen as the most probable solution given this combination of leptons and jets.

It should be noted that this is another important difference in methodology between this analysis and the Run 1 analysis done by Karr and Sliwa[7]. In that analysis, the 2-D smeared jet energy grid was integrated over to get the total likelihood for the event instead of picking the best combination out of all possible smeared jet combinations.

This total likelihood and its corresponding probability distribution will be compared to other combinations of leptons and jets (Bottom Quarks) from the same event where the greater one is favored.

1.4 MET Probability

There is an additional probability factor calculated for each combination based on the measured missing transverse energy. During the analysis of each combination, the energy and momentum of the neutrinos are obtained. The difference between the two missing neutrino x and y momenta in the transverse plane and the missing transverse energy x and y components are compared and assigned a probability based on the shape of the reconstructed $\not\!\!E_T$ x and y components from MC events. This factor is only used to help choose between combinations, it is not used to calculate the mass probability distribution.

1.5 The Joint Likelihood

Once a combination from an event is chosen via the method described above, and all events have been analyzed, a joint likelihood is created. The joint likelihood is the product of the probability distributions from the chosen combinations from each event. Since the true top mass for all of the events is the same, the joint probability distribution reveals where all the events are most consistent in M_t, which should point toward the true mass. The shape of the probability distribution for each event will not necessarily be the same. Some will be asymmetrical and some may have more than one peak. Choosing the mean or the peak value as the top mass for an individual event may include biases from the shape of the distribution. By taking the product of the event distributions, the joint distribution becomes more Gaussian in shape. Multiple peaks and asymmetries in the individual events are eliminated in the joint distribution as are the possible biases that these characteristics may produce. The arithmetic mean of the joint distribution is the top quark mass, the most likely value of the top quark Mass consistent with all of the events in the sample.

1.6 Additional Ways to Choose Combinations

Aside from simply multiplying the most likely combinations together, there are three additional ways of choosing which combinations from an event are used. The method described up to this point is referred to as the "Likelihood" method as it just picks according to the likelihood. The second method involves first choosing the favored combo by likelihood, then searching for its

"twin" where the lepton and jet assignment is switched. Then both of the mass distributions are added together and that combined distribution along with similarly constructed distributions from other events are used to form the joint probability distribution. This method is dubbed the "Likelihood Added" method. The third way starts by finding the two combinations in an event that involved the two jets with the highest transverse momentum, a.k.a. the two leading jets. Only these two combinations are considered and the one with the highest likelihood is chosen as the favorite combination. This method is called the "2 Leading Jets" method, or 2LJ for short. The final way is to start with the two combinations with the two leading jets and simply add their mass distributions together; the likelihood isn't considered except that at least one of the two combinations has to have a non-zero likelihood. This method is known as the "2LJ Added" method. Of all of these methodologies, the one that results in the lowest combined statistical and systematic errors will be used to make the measurement. That method turned out to be the Likelihood Added method and all of the plots and figures shown here will be from the results of that method. Additionally, b-Tagging of the jets was used to help reduce the various backgrounds that plague the di-lepton channel. Both analyses, with and without b-Tagging, will be shown in parallel.

1.7 Backgrounds and MC Generators

The Backgrounds that are expected in the di-leptomn channel are Drell-Yan ($Z\gamma^* \to \tau\tau$ and $Z\gamma^* \to ee/\mu\mu$), Diboson (WW, WZ, and $ZZ \to ll$), $W\gamma \to ll$(without b-Tagging only), and fake evetns. The Drell-Yan and Diboson samples were generated with Pythia Monte Carlo, the $W\gamma$ sample was generated with Baur Monte Carlo, and the fake sample were taken from the real data stream where a jet has been misidentified as a lepton. The signal MC evetns were also generated with Pythia. Table 1 shows a breakdown of the observed number of events by number of b-Tags and lepton pairing. Tables 2 and 3 show the number of events for each each type of background and the signal that are expecte in the real data.

CDF Run II Preliminary (2.0fb^{-1})

# b-Tags	ee	$\mu\mu$	eμ	ll
n = 0	21	18	42	81
n = 1	5	16	24	45
n = 2	4	8	7	19
Events Found in Data (n > 1)	9	24	31	64
Events Found in Data (n > 0)	30	42	73	145

Table 1: Breakdown of the expected number of events by number of b-Tags and lepton pair type.

CDF Run II Preliminary (2.0fb^{-1})

Source	Number of Events
Fakes	21.75 ± 6.33
DY $\rightarrow ee + \mu\mu$	12.78 ± 2.17
WW	6.81 ± 1.17
DY $\rightarrow \tau\tau$	5.26 ± 1.02
WZ	1.59 ± 0.26
ZZ	1.09 ± 0.85
Wγ	0.17 ± 0.18
Total Background	49.45 ± 7.83
$t\bar{t}$ ($\sigma = 6.7$ pb)	93.86 ± 7.14
Total SM Expectation	143.31 ± 13.09
Events Found in Data	145

Table 2: Expected number of events for signal and background in the di-lepton channel for 2.0fb^{-1} with b-Tagging

CDF Run II Preliminary (2.0fb^{-1})

Source	Number of Events
Fakes	2.37 ± 0.69
DY $\rightarrow ee + \mu\mu$	0.84 ± 0.14
WW	0.23 ± 0.04
DY $\rightarrow \tau\tau$	1.04 ± 0.20
WZ	0.04 ± 0.01
ZZ	0.09 ± 0.07
Total Background	4.61 ± 0.74
$t\bar{t}$ ($\sigma = 6.7$ pb)	54.66 ± 4.16
Total SM Expectation	59.27 ± 4.22
Events Observed in Data	64

Table 3: Expected number of events for signal and background in the di-lepton channel for 2.0fb^{-1}

1.8 The Construction of the Pseudo-Experiments

The real data and the Monte Carlo data are treated in exactly the same way in every detail of this analysis except for the method of constructing the joint likelihood. When analyzing real data, the likelihood distributions from all of the candidate events take part in the product that forms the Joint Likelihood. When analyzing the Monte Carlo data, all of the events that pass the cuts and have a solution are put into a pool of events. Next, events are randomly selected from the pool and their likelihood distributions are multiplied together to produce the joint likelihood, which forms one pseudo-experiment. The number of events selected from the pool to form one pseudo-experiment depends on the number of candidate events expected in the real data. The number of events is randomly generated based on a Poisson distribution whose mean is the number of candidate events. Care is taken to not allow an event to appear more than once in the same pseudo-experiment; however, all events in the pool are available to every pseudo-experiment. The mean and RMS from each pseudo-experiment are used to calculate the delta and the pull. The mean, delta, pull, and RMS are each put into histograms to produce their own distributions.

1.9 The RMS Correction

When building the pseudo-experiments using the joint likelihood method described above, it becomes necessary to make a correction to the RMS in order to relate it to the errors. If a simple distribution populated by the means of the events in a pseudo-experiment was used to make the measurement of the mass, the error on that measurement would be related to the RMS of the distribution by $1/\sqrt{N}$. This is, however, not the case for a joint likelihood distribution. A joint likelihood distribution is too narrow and a correction factor is needed.

In order to explain how the correction factor is calculated, a simple model is needed. The model is a box filled with cards, and on each card is a number. The number of cards in the box is N. The sum of all the numbers on the cards is S_{box}, the average of all the numbers is A_{box} and the standard deviation of the numbers in the box is σ_{box}. Now, n cards are drawn from the box at random without replacement. The expectation value for the sum of the draws, S^{EV}_{draw}, is $n * A_{box}$ and the standard error on the sum of the draws from the box is $\sqrt{n} * \sigma_{box}$. The expectation value for the average of the draws, A^{EV}_{draw}, is S^{EV}_{draw}/n which is simply A_{box} and the standard error for the average of the draws is $\sqrt{n} * \sigma_{box}/n$ which simplifies to $\sigma_{box}/\sqrt{n}$.

The pool of events in this analysis is very similar to the box of cards. Each event is a card and the number on the card is the mean of the distribution that corresponds to the event. All of the same statistics that applied to the box of cards will apply to the pool of events in exactly the same way. However, in this analysis, the average is not taken, but instead a joint distribution is made. While at first glance this seems to be very different, making a joint distribution out of the individual mass distributions for each event is like taking the average of the numbers on the cards. So, the standard error for a joint distribution should be similar to that of the random draws from a box and the RMS of a joint distribution is related to that error.

$$RMS_{JD} \propto \sigma_{pool}/\sqrt{n} \tag{19}$$

If all possible combinations of n events were drawn from the pool and the mean of each sample was put into a histogram, that histogram would be a total probability distribution. It would be normal in shape even if the parent distribution is not. Its mean would be similar to the mean of the parent distribution and its width would be the error on the mean of the sample distributions. Since the number of possible combinations is an extremely large number, it is not practical to compute every possible combination. Instead, it is sufficient to draw n events at random from the pool X number of times as long as X is large. When this is done, the resulting histogram will not be a complete total probability histogram, however its mean and width, σ_{XPE}, will be the same as if all possible combinations were sampled. Since the widths of these distributions are the same, σ_{XPE} describes the error on the PEs.

$$\sigma_{XPE} \approx \sigma_{pool}/\sqrt{n} \tag{20}$$

It follows that

$$RMS_{JD} \propto \sigma_{XPE} \tag{21}$$

This gives the correction factor, C_{rms}, to the RMS of the joint distributions.

$$C_{rms} = \sigma_{XPE}/RMS_{JD} \tag{22}$$

The corrected width will be

$$RMS_{corr} = C_{rms} \times RMS_{raw} \tag{23}$$

where RMS_{corr} and RMS_{raw} are the corrected and uncorrected widths of the individual joint distributions of each PE.

Figures 7 and 8 show that C_{RMS} is independent of the top mass.

The value of C_{rms} has a dependence, discussed in more detail in Section 1.11, on the shapes of the likelihood distributions from the events.

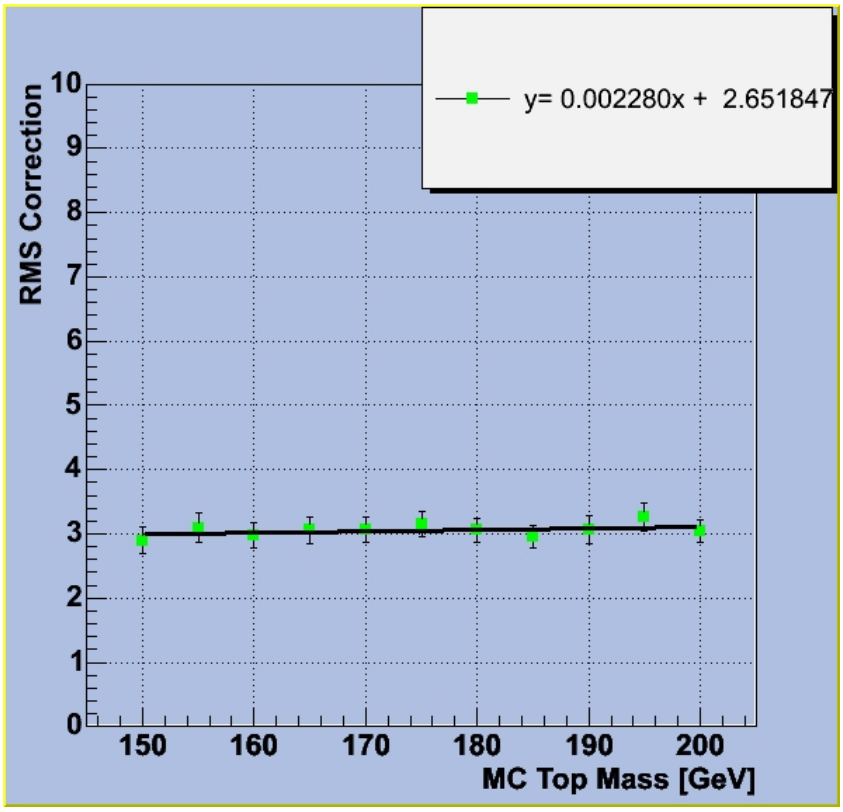

Figure 7: RMS correction factor as a function of top mass for signal-only events.

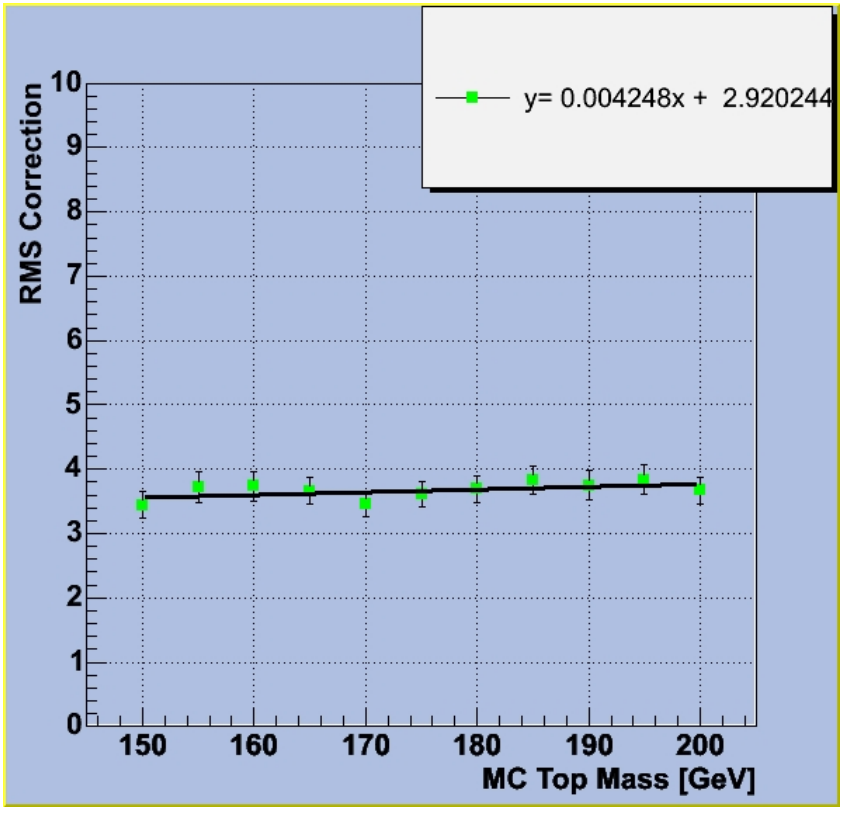

Figure 8: RMS correction factor as a function of top mass for signal and background events.

1.10 The Mapping Function

Analyses of a range of top mass MC samples are used to study the correlation between the generated and reconstructed mass. Given the mean distribution for each mass MC sample, a plot of reconstructed mass as a function of MC mass is made. While it is possible to fit various shapes to the points, a line does the job very well. That fit is the mapping function that will be used to make the final correction to the measurement.

To begin to test the method, parton information from the MC generator level objects are fed into it. This will test the method to see how well it does with the best possible data. Ideally, given that actual 4-momenta of the leptons and b-quarks are being used instead of the reconstructed leptons and jets, if a mass MC sample with M_t of 175 GeV is given to the method, it should return the same mass as its answer. After each of the different mass MC samples are analyzed in this way, the mapping function should come out with a slope of approximately one. Figure 9 shows these mapping function. It is important to note that the event selection was still done on the reconstructed quantities. After the events were selected, only then were the generator level quantities used.

The next step is to use detector level quantities as the input to the analysis. This data will be fully simulated to appear as though it came from the detector, however it will only include $t\bar{t}$ events. It will be more realistic than the generator level data, but will still be more idealistic than real data. Figure 10 shows the resulting mapping functions.

The final step is to add in events that were simulated to look like the various types of background that are expect in the data stream. While this data is simulated, it should look like the real data. Applying the mapping function to the mean of the joint distribution will give the corrected mass of the top quark. The slope of the mapping function is also used to make one of two corrections to the width of the joint distribution to calculate the statistical error. Figure 11 shows the two mapping functions that were used in this analysis. Equations 24 through 31 shows the 8 mapping functions that correspond to the 4 methods of choosing the favored combination.

$$M_t^{raw} = 0.665 \times M_t^{corr} + 50.02 \quad \text{(Likelihood Only)} \tag{24}$$

$$M_t^{raw} = 0.635 \times M_t^{corr} + 53.95 \quad \text{(Likelihood Added)} \tag{25}$$

$$M_t^{raw} = 0.654 \times M_t^{corr} + 52.41 \quad \text{(2 Leading Jets)} \tag{26}$$

$$M_t^{raw} = 0.627 \times M_t^{corr} + 55.85 \quad \text{(2 Leading Jets Added)} \tag{27}$$

$$M_t^{raw} = 0.750 \times M_t^{corr} + 36.364 \quad \text{(Likelihood Only w/ } b\text{-Tagging)} \tag{28}$$

$$M_t^{raw} = 0.747 \times M_t^{corr} + 35.592 \quad \text{(Likelihood Added w/ } b\text{-Tagging)} \tag{29}$$

$$M_t^{raw} = 0.752 \times M_t^{corr} + 36.490 \quad \text{(2 Leading Jets w/ } b\text{-Tagging)} \tag{30}$$

$$M_t^{raw} = 0.754 \times M_t^{corr} + 34.960 \quad \text{(2 Leading Jets Added w/ } b\text{-Tagging)} \tag{31}$$

CDF Run II Preliminary (2.0fb^{-1})

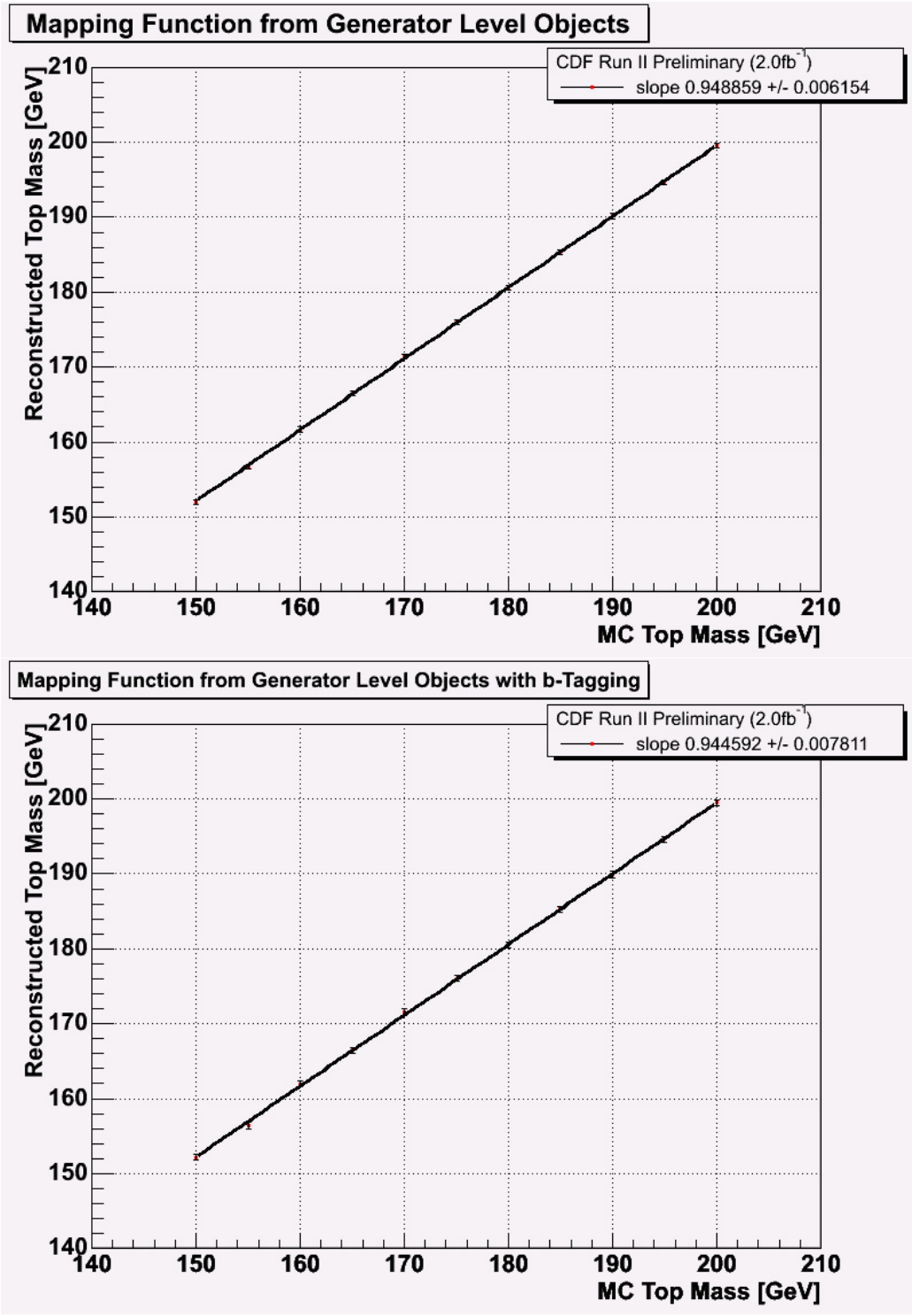

Figure 9: Mapping function from signal-only MC generator level objects.

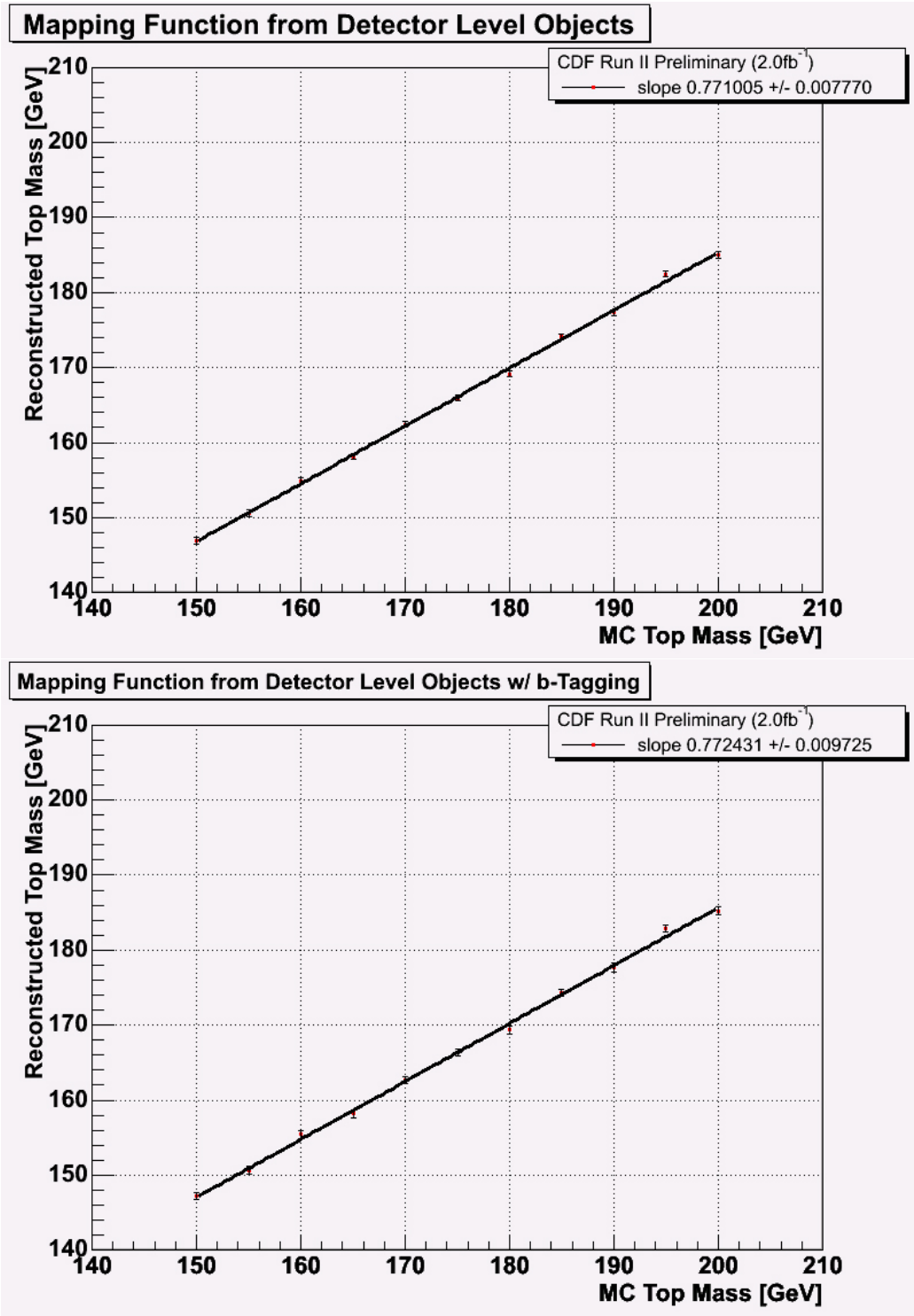

Figure 10: Mapping function from signal-only MC detector level objects.

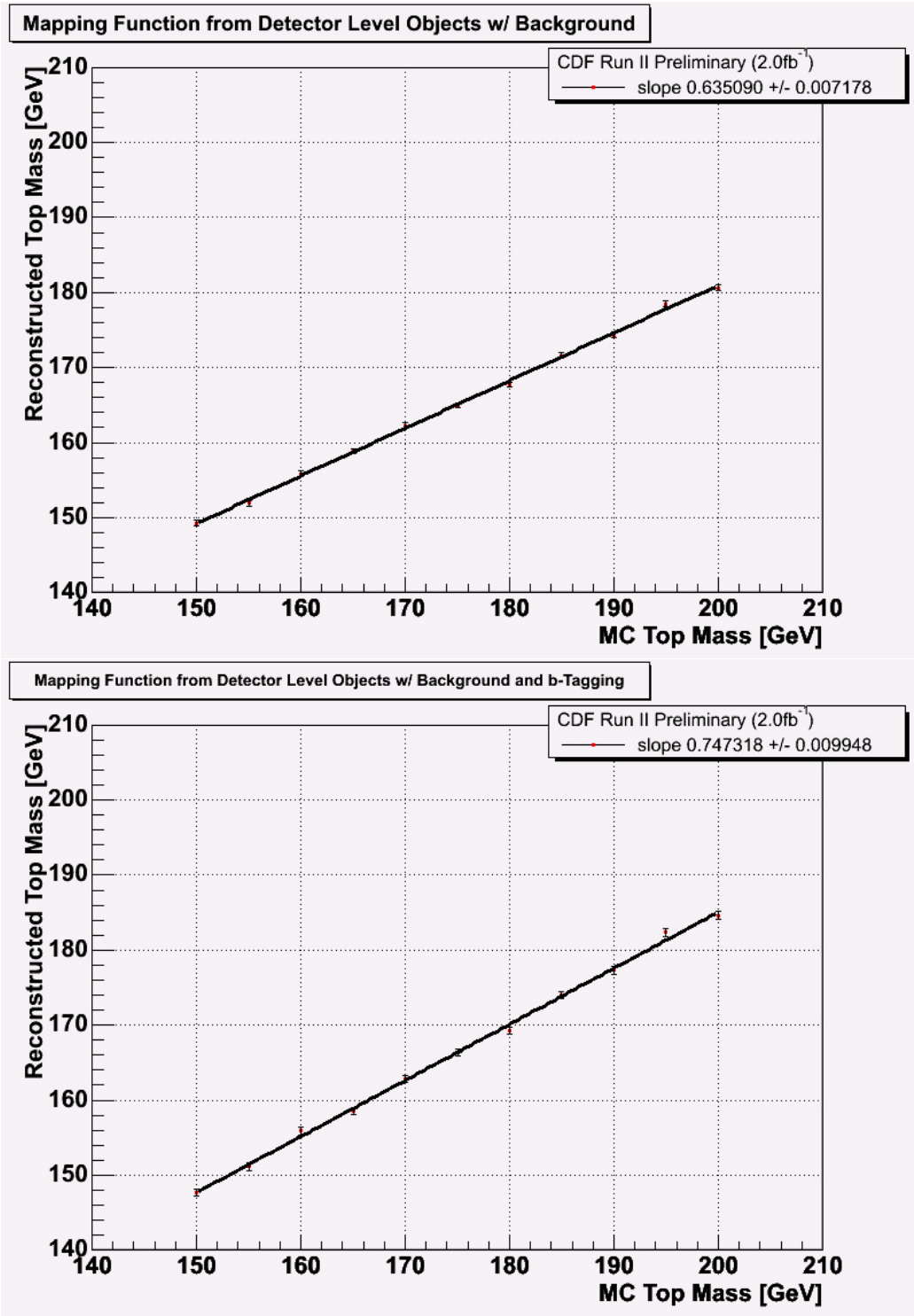

Figure 11: Mapping function from signal and background MC detector level objects.

1.11 Statistical Error

The statistical error on the measurement of the top quark mass is expected to be related to the width of the joint distribution created by the product of the mass distributions from the real data events. In addition to the correction to the RMS as described in Section 1.9, the width of the joint distribution needs to be corrected using the mapping function for the top mass correction. This correction is simply the slope of the mapping function, m_{map}. The statistical error on the measurement of the top quark mass, σ_{stat}^{corr} is related to the raw width of the joint distribution, σ_{JD}^{raw} by:

$$\sigma_{stat}^{corr} = C_{rms} \times \sigma_{JD}^{raw} / m_{map} \tag{32}$$

where C_{rms} is the RMS correction factor. The slopes and RMS correction factors are listed in Table 4.

Combination Choice	Slope	RMS Correction
without b-Tagging		
Likelihood Only	0.665	3.66
Likelihood Added	0.635	3.01
2 Leading Jets	0.654	3.61
2 Leading Jets Added	0.627	3.01
with b-Tagging		
Likelihood Only	0.750	2.77
Likelihood Added	0.747	2.27
2 Leading Jets	0.752	2.73
2 Leading Jets Added	0.754	2.22

Table 4: Correction factors to the width of the top mass joint distribution.

Figure 12 shows all of the correction factors from each data sample, and each method of choosing a combination from an event. The value of the correction factor appears to be related to the shape of the individual event probability distributions and the variation of the means of the individual events. The probability distributions that result from the truth table information tend to be much more symmetrical and smoother, somewhat more Gaussian in shape, than events that use the reconstructed quantities. Additionally, the spread of the means of a given sample is lower when using the truth table than the reconstructed quantities. Together, these effects create the need for a larger correction factor. Its important to note that no matter which type of data is used, the correction factors do not show a significant dependence on the mass of the top quark (See Figures 7 and 8).

Figure 13 shows the distributions of errors that can be expected from each of the strategies of combining jets with leptons after corrections in this analysis based on the corrected widths of individual PEs. These distributions are based on MC data with M_t set to 175 GeV.

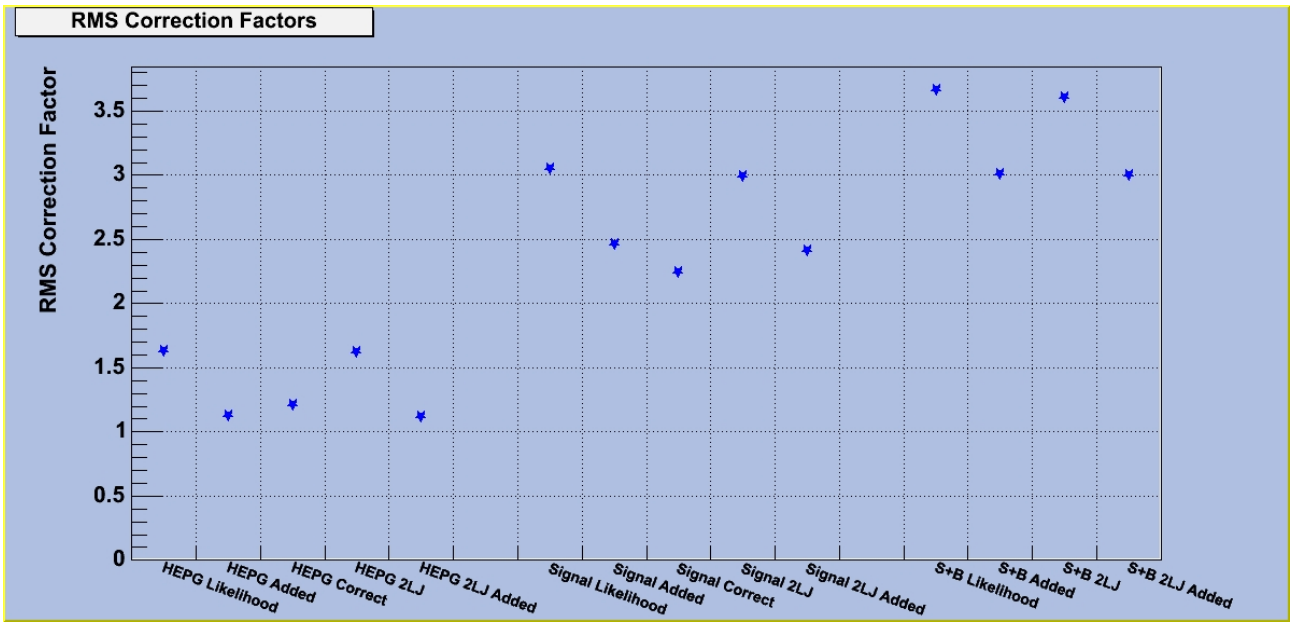

Figure 12: The various correction factors to the RMS by input sample and methodology.

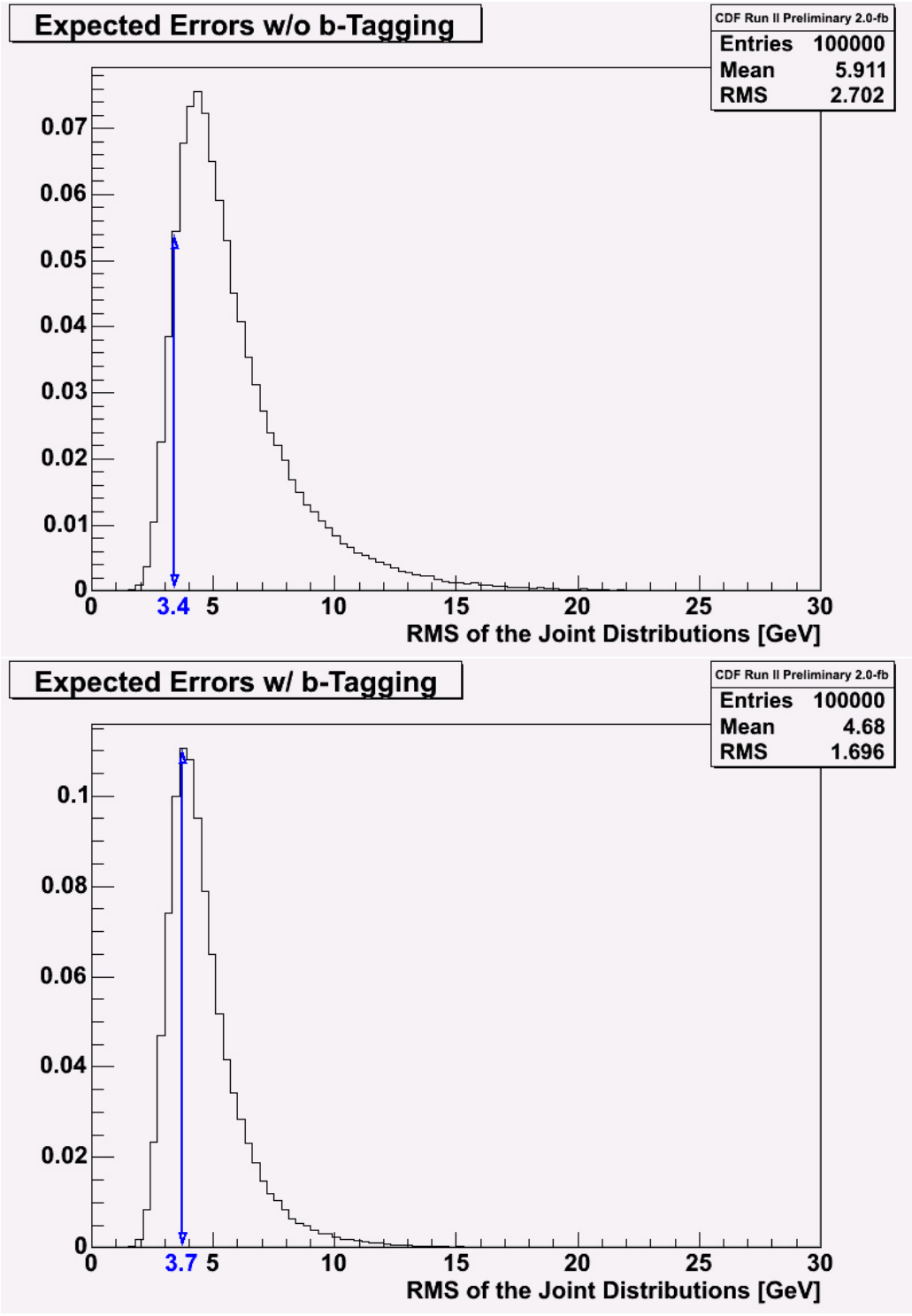

Figure 13: Distributions of corrected widths from the PEs for the Likelihood Added method with b-Tagging

1.12 Sanity Checks

Before it is time to analyze the real data, it is important to have a closer look at how the method affects various variables. If a variable from the signal and background MC data does not match well with the variable in the real data, this would suggest that there is a problem in the method or that the variable is not being modeled well in the MC. The variables of interest are shown in Figures 15 through 18 and include such things like the P_t, η, and ϕ for each of the top quarks, the $t\bar{t}$ -system, and each of the jets and leptons. Figures 15 through 18 show the variables most important to this analysis.

1.13 Systematic Errors

There are nine sources of systematic error that are considered in this analysis. The dominate source comes for the jet energy corrections. One of the strong points of this method is that the Jet Energy Scale systematic uncertanty is small as compared to other analyses, which gives the lowest total systematic uncertanty of any analysis done in the di-lepton channel at CDF and DØ given a comparable dataset. Many of the remaining uncertanties, those with a value of 0.33 GeV and 0.42 GeV, showed an insignificant shift when compared to the MC statistical error. In these cases, the value quoted was the limiting MC statistical error. Table 5 shows the nine sources of systematic error and thier values.

CDF Run II Preliminary (2.0fb^{-1})

Likelihood Added		
Source	Without b-Tagging	With b-Tagging
Jet Energy Scale	± 1.42 GeV	± 0.97 GeV
PDF	± 0.74 GeV	± 0.31 GeV
ISR/FSR	± 0.46 GeV	± 0.42 GeV
MC Generator	± 0.33 GeV	± 0.42 GeV
Background	± 0.35 GeV	± 0.42 GeV
b-Jet Energy Scale	± 0.33 GeV	± 0.42 GeV
Lepton P_T	± 0.33 GeV	± 0.42 GeV
Pile-up	± 0.93 GeV	± 0.72 GeV
Color Reconnection	± 0.41 GeV	± 0.54 GeV
Total Systematic Uncertainty	± 2.06 GeV	± 1.65 GeV

Table 5: The systematic uncertainties in the di-lepton channel.

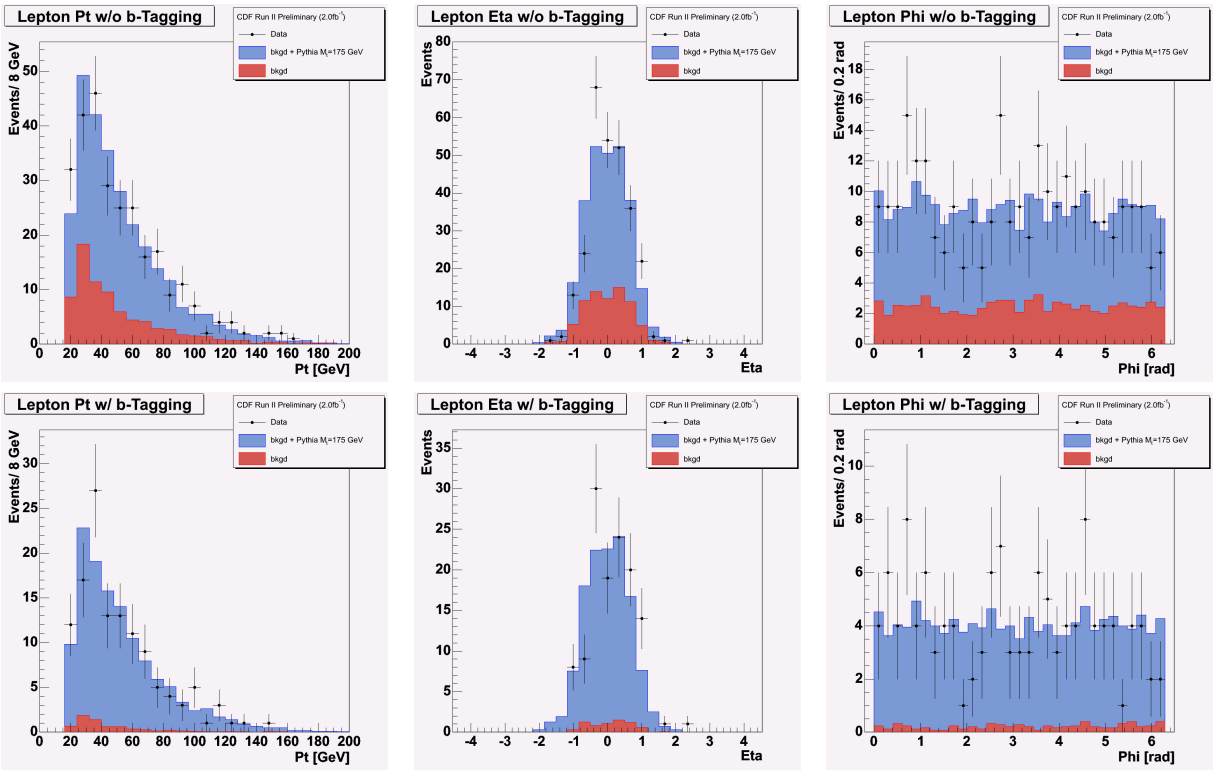

Figure 14: The P_t, η, and ϕ for the leptons in both the MC data and the real data.

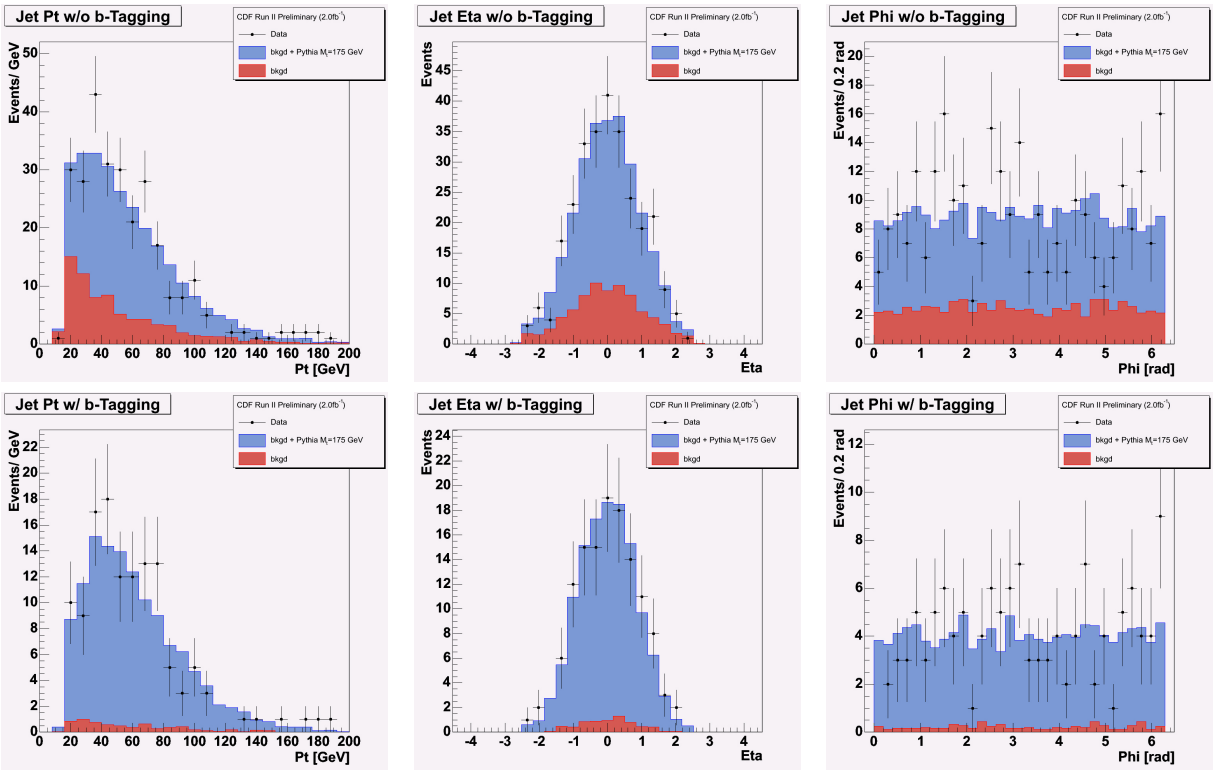

Figure 15: The P_t, η, and ϕ for the jets in both the MC data and the real data.

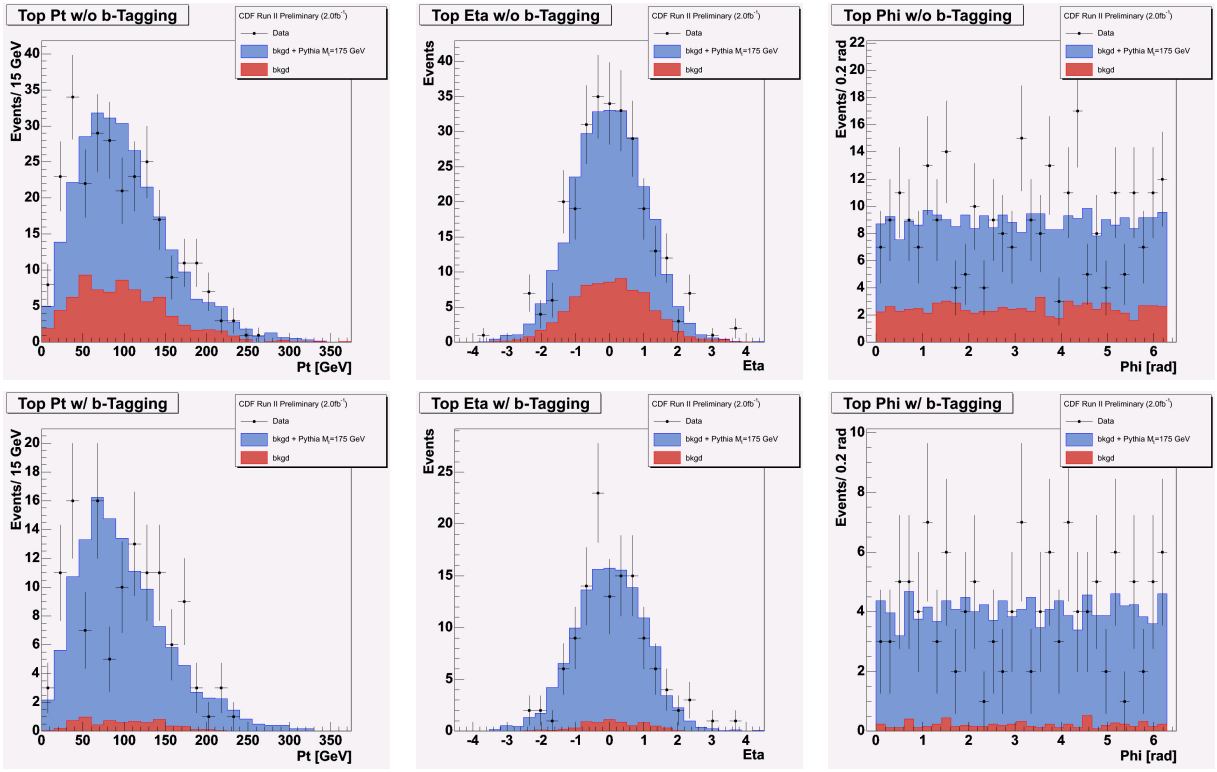

Figure 16: The P_t, η, and ϕ for the top quarks in both the MC data and the real data.

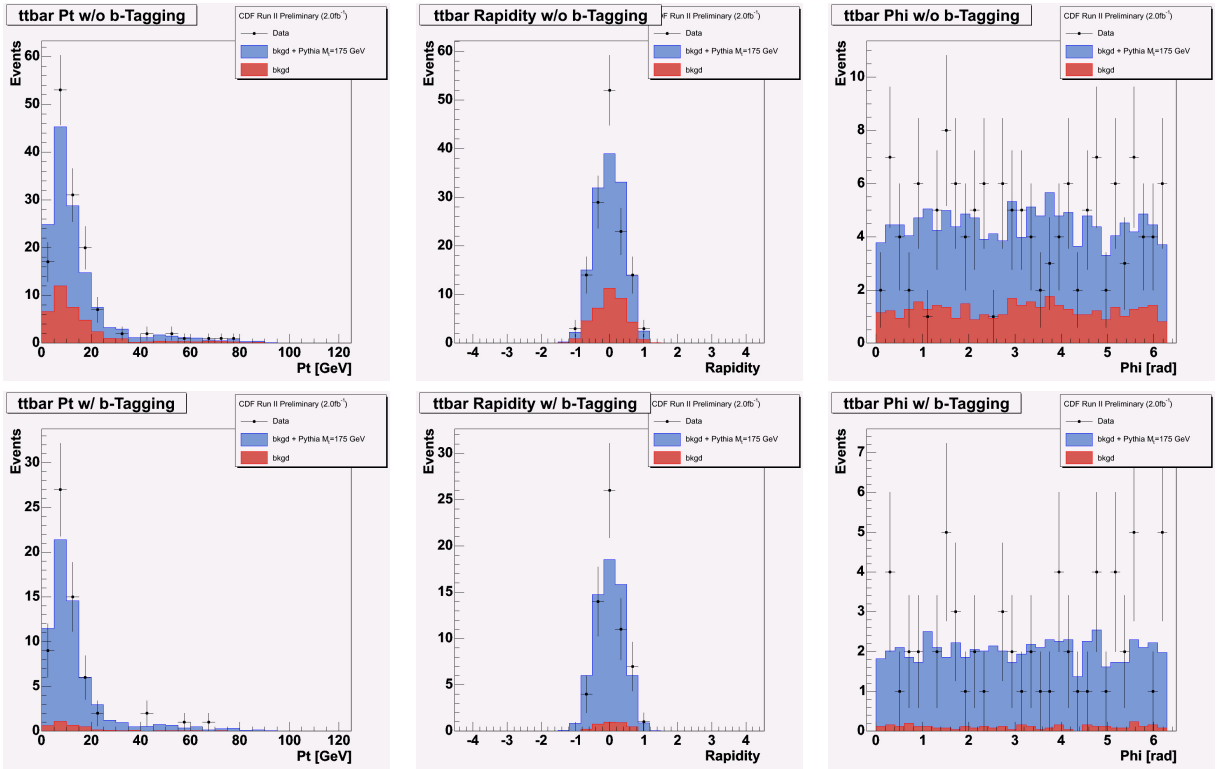

Figure 17: The P_t, rapidity, and ϕ for the top-anti-top quark system in both the MC data and the real data.

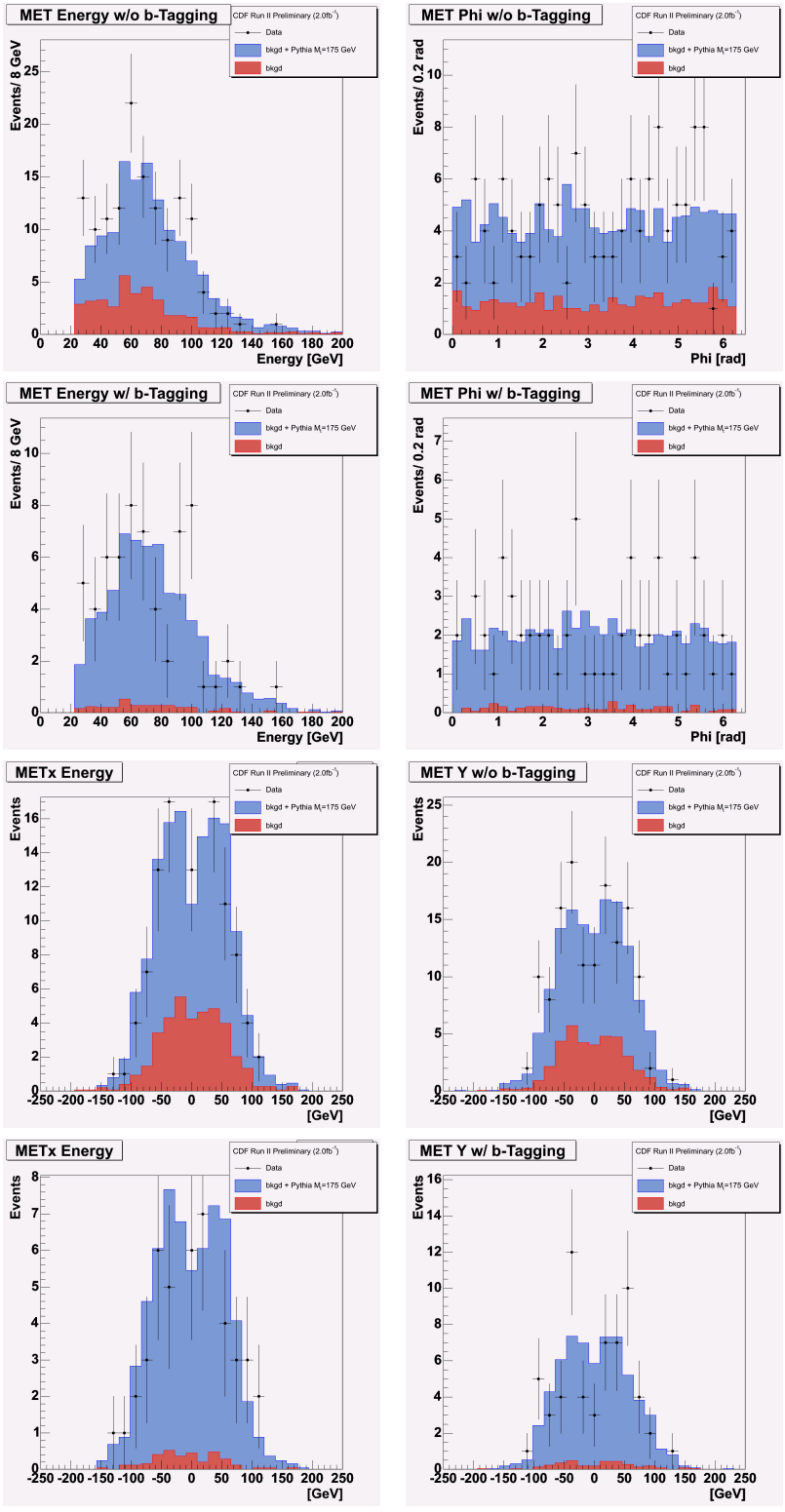

Figure 18: The energy, ϕ, energy in the x-direction, and energy in the y-direction for the $\not\!\!E_T$ in both the MC data and the real data.

1.14 Results

The Dalitz-Goldstein top mass fitting technique has been applied to the di-lepton events in Run II of the CDF experiment and the top quark mass has been determined to be:

$$M_{top} = 172.3 \pm 3.4(stat.) \pm 2.1(syst.) GeV/c^2. \quad (33)$$

$$M_{top} = 170.5 \pm 3.7(stat.) \pm 1.7(syst.) GeV/c^2. \quad (34)$$

Table 6 shows the uncorrected (raw) results for each of the four ways of choosing a combination from an event with and without b-tagging. Using Eqs. 24 through 31, the raw masses can be corrected. To correct the raw RMSs, Eq. 32 and the values from Table 4 were used. Table 7 shows the fully corrected results.

Method	Mass	Statistical Error
without b-tagging		
Likelihood Only	163.6	±0.76
Likelihood Added	163.3	±0.72
2 Leading Jets	162.7	±1.25
2 Leading Jets Added	162.8	±1.04
with b-tagging		
Likelihood Only	164.4	±1.28
Likelihood Added	163.0	±1.23
2 Leading Jets	164.4	±1.24
2 Leading Jets Added	162.9	±1.20

Table 6: Results without any corrections. (Raw)

Method	Mass	Statistical (expected SE)	Systematic	Errors in Quad.
without b-tagging				
Likelihood Only	170.7	±4.20 (4.61)	±1.97	±4.64
Likelihood Added	172.3	±3.41 (4.31)	±2.06	±3.99
2 Leading Jets	168.8	±6.88 (4.59)	±2.07	±7.19
2 Leading Jets Added	170.3	±5.00 (4.44)	±1.99	±5.38
with b-tagging				
Likelihood Only	170.7	±4.73 (4.36)	±2.06	±5.16
Likelihood Added	170.5	±3.74 (3.87)	±1.65	±4.09
2 Leading Jets	170.1	±4.50 (4.78)	±2.43	±5.12
2 Leading Jets Added	169.7	±3.54 (4.40)	±1.69	±3.92

Table 7: Final results with corrections.

The results shown in Eqs. 33 and 34 have been chosen to be the the main results from this analysis based on the facts that they use all jets in an event and consistently have low systematic errors. They are therefore considered to be the best measurements from this method.

References

[1] R. H. Dalitz and G. Goldstein. Decay and Polarization Properties of the Top Quark. *Phys. Rev. D*, 45:1531, 1992.

[2] R.H.Dalitz G. Goldstein and K. Sliwa. On Observing the Top Quark Production at the Tevatron. CDF Note 1750, CDF, 1992.

[3] R.H.Dalitz G. Goldstein and K. Sliwa. Search for the Top-Antitop Events in the Semileptonic Mode. CDF Note 1751, CDF, 1992.

[4] R.H.Dalitz G. Goldstein and K. Sliwa. Observing the Top Quark Production at the Fermilab Tevatron. *Phys. Rev. D*, 47:967, 1993.

[5] K. Sliwa. Search for the ttbar Events in the Lepton and Four Jets Final State. CDF Note 1993, CDF, 1993.

[6] K. Sliwa. A Non Black Box Approach to Kinematical Fitting and Its Application to Top Quark Search. CDF Note 3057, CDF, 1995.

[7] Kristo M. Karr. *Measurement of the Top Quark Mass by Application of the Dalitz-Goldstein Method to Dilepton Events*. PhD thesis, Tufts University, 1999.